

PAGE: 1

**RAW SEQUENCE LISTING
PATENT APPLICATION US/08/785,455**DATE: 02/25/97
TIME: 11:24:02**INPUT SET: S15668.raw**

This Raw Listing contains the General Information Section and up to the first 5 pages

ENTERED

1 **SEQUENCE LISTING**2 **3 (1) General Information**

4

5 **(i) APPLICANT**

- 6 (A) NAME: Hodgson, John
7 (B) STREET: Great West Road
8 (C) CITY: Brentford
9 (D) STATE OR PROVINCE: Middlesex
10 (E) COUNTRY: United Kingdom
11 (F) POSTAL CODE: TW8 9EP

12

13

- 14 (A) NAME: Hodgson, John
15 (B) STREET: Great West Road
16 (C) CITY: Brentford
17 (D) STATE OR PROVINCE: Middlesex
18 (E) COUNTRY: United Kingdom
19 (F) POSTAL CODE: TW8 9EP

20

21 (ii) TITLE OF THE INVENTION: Novel tRNA Synthetase

22

23

24 (iii) NUMBER OF SEQUENCES: 2

25

26 (iv) COMPUTER-READABLE FORM:

- 27 (A) MEDIUM TYPE: Diskette
28 (B) COMPUTER: IBM Compatible
29 (C) OPERATING SYSTEM: DOS
30 (D) SOFTWARE: FastSEQ for Windows Version 2.0

31

32 (v) CURRENT APPLICATION DATA:

- 33 (A) APPLICATION NUMBER:
34 (B) FILING DATE:

35

36 (2) INFORMATION FOR SEQ ID NO:1:

37

38 (i) SEQUENCE CHARACTERISTICS:

- 39 (A) LENGTH: 1974 base pairs
40 (B) TYPE: nucleic acid
41 (C) STRANDEDNESS: double
42 (D) TOPOLOGY: linear

43

44 (ii) MOLECULE TYPE: Genomic DNA

45

46 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/785,455DATE: 02/25/97
TIME: 11:24:07

INPUT SET: S15668.raw

47 ATGGCTAAAG AAACATTTA TATAACAACC CCAATATACT ATCCTAGTGG GAATTACAT 60
 48 ATAGGACATG CATATTCTAC AGTGGCTGGA GATGTTATTG CAAGATATAA GAGAATGCAA 120
 50 GGATATGATG TTCGTTATTG GACTGGAACG GATGAACACG GTCAAAAAAT TCAAGAAAAA 180
 51 GCTAAAAAG CTGGTAAGAC AGAAAATTGAA TATTTGGATG AGATGATTGC TGGAATTAAA 240
 52 CAATTGTGGG CTAAGCTTGA AATTCAAAT GATGATTTTA TCAGAACAAAC TGAAGAACGT 300
 53 CATAAACATG TCGTTGAGCA AGTGTGAA CGTTTATTAA AGCAAGGTGA TATCTATTAA 360
 54 GGTGAATATG AAGGTTGGTA TTCTGTTCCG GATGAAACAT ACTATACAGA GTCACAATTA 420
 55 GTAGACCCAC AATACGAAAA CGGTAAGAA ATTGGTGGCA AAAGTCCAGA TTCTGGACAC 480
 56 GAAGTTGAAC TAGTTAAAGA AGAAAAGTTAT TTCTTTAATA TTAGTAAATA TACAGACCGT 540
 57 TTATTAGAGT TCTATGACCA AAATCCAGAT TTTATACAAAC CACCATCAAG AAAAAATGAA 600
 58 ATGATTAACA ACTTCATTAA ACCAGGACTT GCTGATTTAG CTGTTCTCG TACATCATT 660
 59 AACTGGGTG TCCATGTTCC GTCTAATCCA AAACATGTTG TTTATGTTG GATTGATGCG 720
 60 TTAGTTAACT ATATTCAGC ATTAGGCTAT TTATCAGATG ATGAGTCACT ATTTAACAAA 780
 61 TACTGCCAG CAGATATTCA TTTAATGGCT AAGGAAATTG TGGCATTCCA CTCAATTATT 840
 62 TGGCCTATT TATTGATGGC ATTAGACTTA CCGTTACCTA AAAAAGTCTT TGCACATGGT 900
 63 TGGATTTGA TGAAAGATGG AAAAATGAGT AAATCTAAAG GTAATGTTGT AGACCCTAAT 960
 64 ATTTTAATTG ATCGCTATGG TTTAGATGCT ACACGTTATT ATCTAATGCG TGAATTACCA 1020
 65 TTTGGTTCAG ATGGCGTATT TACACCTGAA GCATTTGTTG AGCGTACAAA TTTCGATCTA 1080
 66 GCAAATGACT TAGGTAACCT AGTAAACCGT ACGATTCTA TGGTTAATAA GTACTTTGAT 1140
 67 GGCGAATTAC CAGCGTATCA AGGTCCACTT CATGAATTAG ATGAAGAAAT GGAAGCTATG 1200
 68 GCTTAGAAA CAGTGAAAAG CTACACTGAA AGCATGGAAA GTTGCAATT TTCTGTGGCA 1260
 69 TTATCTACGG TATGGAAGTT TATAAGTAGA ACGAATAAGT ATATTGACGA ACAACGCC 1320
 70 TGGGTATTAG CTAAGGACGA TAGCCAAAAA GATATGTTAG GCAATGTAAT GGCTCACTTA 1380
 71 GTTAAAATA TTCGTTATGC AGCTGTATTA TTACGTCCAT TCTTAACACA TGCGCCGAAA 1440
 72 GAGATTTTG AACAAATTGAA CATAAAACAA CCTCAATTAA TTGAAATTAG TAGTTAGAG 1500
 73 CAATATGGT TGCTTACTGA GTCAATTATC GTTACTGGGC AACCTAAACC TATTTCCCA 1560
 74 AGATTGGATA GCGAAGCGGA AATTGCATAT ATCAAAGAAAT CAATGCAACC GCCTGCTACT 1620
 75 GAAGAGGAAA AAGAAGAGAT TCCTAGCAAA CCTCAAATTG ATATTAAAGA CTTTGATAAA 1680
 76 GTTGAATTAA AGGCAGCAAC GATTATTGAT GCTGAACATG TTAAGAAGTC AGATAAGCTT 1740
 77 TTAAAAATTG AAGTAGACTT AGATTCTGAA CAAAGACAAA TTGTATCAGG AATTGCCAAA 1800
 78 TTCTATACAC CAGATGATAT TATTGGTAA AAAGTAGCAG TTGTTACTAA CCTGAAACCA 1860
 79 GCTAAATTAA TGGGACAAAA ATCTGAAGGT ATGATATTAT CTGCTGAAAA AGATGGTGT 1920
 80 TTAACCTTAG TAAGTTTACCA AAGTGAATT CCAAATGGTG CAGTGATTAA ATAA 1974
 81
 82 (2) INFORMATION FOR SEQ ID NO:2:
 83
 84 (i) SEQUENCE CHARACTERISTICS:
 85 (A) LENGTH: 657 amino acids
 86 (B) TYPE: amino acid
 87 (C) STRANDEDNESS: single
 88 (D) TOPOLOGY: linear
 89
 90 (ii) MOLECULE TYPE: peptide
 91
 92 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
 93
 94 Met Ala Lys Glu Thr Phe Tyr Ile Thr Thr Pro Ile Tyr Tyr Pro Ser
 95 1 5 10 15
 96 Gly Asn Leu His Ile Gly His Ala Tyr Ser Thr Val Ala Gly Asp Val
 97 20 25 30
 98 Ile Ala Arg Tyr Lys Arg Met Gln Gly Tyr Asp Val Arg Tyr Leu Thr
 99 35 40 45

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/785,455DATE: 02/25/97
TIME: 11:24:12

INPUT SET: S15668.raw

100 Gly Thr Asp Glu His Gly Gln Lys Ile Gln Glu Lys Ala Gln Lys Ala
 101 50 55 60
 102 Gly Lys Thr Glu Ile Glu Tyr Leu Asp Glu Met Ile Ala Gly Ile Lys
 103 65 70 75 80
 104 Gln Leu Trp Ala Lys Leu Glu Ile Ser Asn Asp Asp Phe Ile Arg Thr
 105 85 90 95
 106 Thr Glu Glu Arg His Lys His Val Val Glu Gln Val Phe Glu Arg Leu
 107 100 105 110
 108 Leu Lys Gln Gly Asp Ile Tyr Leu Gly Glu Tyr Glu Gly Trp Tyr Ser
 109 115 120 125
 110 Val Pro Asp Glu Thr Tyr Thr Glu Ser Gln Leu Val Asp Pro Gln
 111 130 135 140
 112 Tyr Glu Asn Gly Lys Ile Ile Gly Gly Lys Ser Pro Asp Ser Gly His
 113 145 150 155 160
 114 Glu Val Glu Leu Val Lys Glu Glu Ser Tyr Phe Phe Asn Ile Ser Lys
 115 165 170 175
 116 Tyr Thr Asp Arg Leu Leu Glu Phe Tyr Asp Gln Asn Pro Asp Phe Ile
 117 180 185 190
 118 Gln Pro Pro Ser Arg Lys Asn Glu Met Ile Asn Asn Phe Ile Lys Pro
 119 195 200 205
 120 Gly Leu Ala Asp Leu Ala Val Ser Arg Thr Ser Phe Asn Trp Gly Val
 121 210 215 220
 122 His Val Pro Ser Asn Pro Lys His Val Val Tyr Val Trp Ile Asp Ala
 123 225 230 235 240
 124 Leu Val Asn Tyr Ile Ser Ala Leu Gly Tyr Leu Ser Asp Asp Glu Ser
 125 245 250 255
 126 Leu Phe Asn Lys Tyr Trp Pro Ala Asp Ile His Leu Met Ala Lys Glu
 127 260 265 270
 128 Ile Val Arg Phe His Ser Ile Ile Trp Pro Ile Leu Leu Met Ala Leu
 129 275 280 285
 130 Asp Leu Pro Leu Pro Lys Lys Val Phe Ala His Gly Trp Ile Leu Met
 131 290 295 300
 132 Lys Asp Gly Lys Met Ser Lys Ser Lys Gly Asn Val Val Asp Pro Asn
 133 305 310 315 320
 134 Ile Leu Ile Asp Arg Tyr Gly Leu Asp Ala Thr Arg Tyr Tyr Leu Met
 135 325 330 335
 136 Arg Glu Leu Pro Phe Gly Ser Asp Gly Val Phe Thr Pro Glu Ala Phe
 137 340 345 350
 138 Val Glu Arg Thr Asn Phe Asp Leu Ala Asn Asp Leu Gly Asn Leu Val
 139 355 360 365
 140 Asn Arg Thr Ile Ser Met Val Asn Lys Tyr Phe Asp Gly Glu Leu Pro
 141 370 375 380
 142 Ala Tyr Gln Gly Pro Leu His Glu Leu Asp Glu Glu Met Glu Ala Met
 143 385 390 395 400
 144 Ala Leu Glu Thr Val Lys Ser Tyr Thr Glu Ser Met Glu Ser Leu Gln
 145 405 410 415
 146 Phe Ser Val Ala Leu Ser Thr Val Trp Lys Phe Ile Ser Arg Thr Asn
 147 420 425 430
 148 Lys Tyr Ile Asp Glu Thr Thr Pro Trp Val Leu Ala Lys Asp Asp Ser
 149 435 440 445
 150 Gln Lys Asp Met Leu Gly Asn Val Met Ala His Leu Val Glu Asn Ile
 151 450 455 460
 152 Arg Tyr Ala Ala Val Leu Leu Arg Pro Phe Leu Thr His Ala Pro Lys

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/785,455DATE: 02/25/97
TIME: 11:24:18

INPUT SET: S15668.raw

153	465	470	475	480
154	Glu Ile Phe Glu Gln Leu Asn Ile Asn Asn Pro Gln Phe Met Glu Phe			
155	485	490	495	
156	Ser Ser Leu Glu Gln Tyr Gly Val Leu Thr Glu Ser Ile Met Val Thr			
157	500	505	510	
158	Gly Gln Pro Lys Pro Ile Phe Pro Arg Leu Asp Ser Glu Ala Glu Ile			
159	515	520	525	
160	Ala Tyr Ile Lys Glu Ser Met Gln Pro Pro Ala Thr Glu Glu Glu Lys			
161	530	535	540	
162	Glu Glu Ile Pro Ser Lys Pro Gln Ile Asp Ile Lys Asp Phe Asp Lys			
163	545	550	555	560
164	Val Glu Ile Lys Ala Ala Thr Ile Ile Asp Ala Glu His Val Lys Lys			
165	565	570	575	
166	Ser Asp Lys Leu Leu Lys Ile Gln Val Asp Leu Asp Ser Glu Gln Arg			
167	580	585	590	
168	Gln Ile Val Ser Gly Ile Ala Lys Phe Tyr Thr Pro Asp Asp Ile Ile			
169	595	600	605	
170	Gly Lys Lys Val Ala Val Val Thr Asn Leu Lys Pro Ala Lys Leu Met			
171	610	615	620	
172	Gly Gln Lys Ser Glu Gly Met Ile Leu Ser Ala Glu Lys Asp Gly Val			
173	625	630	635	640
174	Leu Thr Leu Val Ser Leu Pro Ser Ala Ile Pro Asn Gly Ala Val Ile			
175	645	650	655	
176	Lys			
177				
178				